Available online at www.sciencedirect.com
SCIENCE DIRECT®

Stochastic Processes and their Applications 116 (2006) 779–795

stochastic processes and their applications

www.elsevier.com/locate/spa

Backward stochastic Volterra integral equations and some related problems

Jiongmin Yong*

Department of Mathematics, University of Central Florida, Orlando, FL 32816, United States

Received 14 June 2004; received in revised form 28 October 2004; accepted 16 January 2006
Available online 9 February 2006

Abstract

Backward stochastic Volterra integral equations (BSVIEs, for short) are introduced. The existence and uniqueness of adapted solutions are established. A duality principle between linear BSVIEs and (forward) stochastic Volterra integral equations is obtained. As applications of the duality principle, a comparison theorem is proved for the adapted solutions of BSVIEs, and a Pontryagin type maximum principle is established for an optimal control of stochastic integral equations.

MSC: 60H10

Keywords: Backward stochastic Volterra integral equation; Adapted solutions; Duality principle; Comparison theorem; Pontryagin maximum principle

1. Introduction

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ be a complete filtered probability space, and $W(\cdot)$ be a d-dimensional standard Brownian motion whose natural filtration is given by $\mathbb{F} \overset{\Delta}{=} \{\mathcal{F}_t\}_{t\geq 0}$. By an Itô type (forward) stochastic differential equation (FSDE, for short), we mean the following initial value problem:

$$\begin{cases} dX(t) = b(t, X(t))\, dt + \sigma(t, X(t))\, dW(t), \quad t \in [0, T], \\ X(0) = x. \end{cases} \tag{1.1}$$

* Tel.: +1 407 8232089; fax: +1 407 8236253.
E-mail address: jyong@mail.ucf.edu.

doi:10.1016/j.spa.2006.01.005

Standard results concerning (1.1) can be found in many books (see [6,7], for examples). We know that (1.1) has the following equivalent integral form:

$$X(t) = x + \int_0^t b(s, X(s))\,\mathrm{d}s + \int_0^t \sigma(s, X(s))\,\mathrm{d}W(s), \quad t \in [0, T]. \tag{1.2}$$

This suggests naturally to us that one may consider the following type integral equation:

$$X(t) = f(t) + \int_0^t b(t, s, X(s))\,\mathrm{d}s + \int_0^t \sigma(t, s, X(s))\,\mathrm{d}W(s), \quad t \in [0, T]. \tag{1.3}$$

The above is referred to as a forward stochastic Volterra integral equation (FSVIE, for short) for which some systematic studies (even for much more general cases) were carried out in the literature (see [1,5,13,17,19,22], for examples). It is clear that in general, one cannot transform (1.3) into an FSDE of form (1.1), even if the coefficients b and σ are smooth. Thus, mathematically, (1.3) is strictly more general than (1.2). On the other hand, from a practical applications point of view, (1.3) allows some long-range dependence of the noise in the models under consideration. It is interesting that one could even allow $\sigma(t, s, X(s))$ to be only $\mathcal{F}_t$-measurable in some way (thus, anticipating integrals have to be involved), but one still might have adapted solutions (see [17]). Hence, theory for (1.3) is much richer than that of (1.2), both in theory and applications.

On the other hand, in 1973, while studying stochastic optimal control problems, Bismut introduced (linear) backward stochastic differential equations (BSDEs, for short) for the first time [2]. Pardoux and Peng [16] first studied the general nonlinear BSDEs of the following form in 1990:

$$\begin{cases} \mathrm{d}Y(t) = h(t, Y(t), Z(t))\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t), & t \in [0, T], \\ Y(T) = \xi. \end{cases} \tag{1.4}$$

This is a terminal value problem for an Itô type stochastic differential equation. By an adapted solution, we mean a pair $(Y(\cdot), Z(\cdot))$ of $\{\mathcal{F}_r\}_{r \geq 0}$-adapted processes satisfying (1.4) in the usual Itô sense. Since 1990, there has appeared a large volume of literature published related to the theory and applications for BSDEs ([3,8,9,12,15,18,23] and references cited therein).

Now, let us take the viewpoint from the relation between (1.1) and (1.3). As we know (1.4) is understood as the following integral equation:

$$Y(t) = \xi - \int_t^T h(s, Y(s), Z(s))\,\mathrm{d}s - \int_t^T Z(s)\,\mathrm{d}W(s), \quad t \in [0, T]. \tag{1.5}$$

We can call the above a backward stochastic Volterra integral equation (BSVIE, for short). Inspired by (1.3), we come up with the following natural question:

What is the analog of (1.3) for (1.5) as (1.3) for (1.2)?

It turns out that an analog should take the following form:

$$Y(t) = f(t) - \int_t^T h(t, s, Y(s), Z(t, s), Z(s, t))\,\mathrm{d}s - \int_t^T Z(t, s)\,\mathrm{d}W(s), \quad t \in [0, T], \tag{1.6}$$

where $f(\cdot)$ and $h(\cdot)$ are given, and one is looking for the pair $(Y(\cdot), Z(\cdot\,,\cdot))$. We call (1.6) a BSVIE. Apparently, there are two interesting features of (1.6):

(i) The term $Z(t,s)$ depends on t;
(ii) The drift depends on both $Z(t,s)$ and $Z(s,t)$ in general (we will see the interesting role that such a dependence plays).

By taking conditional expectation on both sides of (1.6), we have

$$Y(t) = E\left[f(t) - \int_t^T h(t,s,Y(s),Z(t,s),Z(s,t))\,\mathrm{d}s \,\middle|\, \mathcal{F}_t \right], \quad t \in [0,T]. \tag{1.7}$$

This is related to another interesting motivation of studying BSVIEs. In 1992, Duffie and Epstein introduced the so-called stochastic differential utility [6], by which we mean a solution to the following equation:

$$Y(t) = E\left[\xi + \int_t^T h(s,Y(s))\,\mathrm{d}s \,\middle|\, \mathcal{F}_t \right], \quad t \in [0,T]. \tag{1.8}$$

Recall that the standard expected utility admits the following representation:

$$Y(t) = E\left[\xi \mathrm{e}^{-\beta(T-t)} + \int_t^T u(C(s))\mathrm{e}^{-\beta(s-t)}\,\mathrm{d}s \,\middle|\, \mathcal{F}_t \right], \quad t \in [0,T], \tag{1.9}$$

where $C(\cdot)$ is a consumption process, $u(\cdot)$ is a utility function and β is the discount rate. It is known that (see [9,14]), (1.8) is equivalent to a BSDE. On the other hand, although the integrand in (1.9) depends on t, due to its special form, it is also equivalent to a BSDE. We note that the factor $\mathrm{e}^{-\beta(s-t)}$ in the integral term of (1.9) exhibits some memory effect. Thus, one can naturally consider more general memory case, which then leads to an equation of form (1.7). It is not hard to see that one is not able to transform (1.7) into a BSDE, in general. In [9,20], optimal control problems were studied for general stochastic differential utilities. Some other relevant results can be found in [4,14,21] and references cited therein. It is expected that the theory of BSVIEs presented in this paper should lead to some further generalization of the results in the above-mentioned works.

In this paper, we will establish a preliminary theory for BSVIEs. Besides the well-posedness of BSVIEs, we will also present a duality principle between linear FSVIEs and linear BSVIEs, a comparison theorem for BSVIEs, and a Pontryagin type maximum principle for an optimal control problem of stochastic integral equations.

The rest of the paper is organized as follows. Section 2 is devoted to establishing some preliminary results. In Section 3, we prove the existence and uniqueness of adapted solutions to general BSVIEs, together with some stability results. In Section 4, we present a duality principle for linear FSVIEs and linear BSVIEs. Finally, in Sections 5 and 6, as applications of the duality principle, we prove a comparison theorem for one-dimensional BSVIEs, and a Pontryagin type maximum principle for optimal control of stochastic integral equations.

2. Backward stochastic integral equations

In this section, we are going to look at a simple BSVIE which will play an interesting role later.

To begin with, let us introduce/recall some spaces. Let $H = \mathbb{R}^m, \mathbb{R}^{m\times d}$, etc., whose norm is denoted by $|\cdot|$. Let $\mathcal{B}(G)$ be the Borel σ-field of metric space G. We define

$$\begin{aligned}
&L^2(\Omega) = \{\xi : \Omega \to H \mid \xi \text{ is } \mathcal{F}_T\text{-measurable, } E|\xi|^2 < \infty\},\\
&L^2((0,T)\times\Omega) = \Big\{\varphi : (0,T)\times\Omega \to H \mid \varphi(\cdot) \text{ is } \mathcal{B}([0,T])\otimes\mathcal{F}_T\text{-measurable},\\
&\qquad E\int_0^T |\varphi(t)|^2\,\mathrm{d}t < \infty\Big\},\\
&\overline{C}([0,T]; L^2(\Omega)) = \Big\{\varphi(\cdot) \in L^2((0,T)\times\Omega) \mid \varphi(t) \text{ is } \mathcal{F}_T\text{-measurable}, \forall t \in [0,T],\\
&\qquad \varphi(\cdot) \text{ is continuous from } [0,T] \text{ to } L^2(\Omega), \sup_{t\in[0,T]} E|\varphi(t)|^2 < \infty\Big\},\\
&C([0,T]; L^2(\Omega)) = \{\varphi(\cdot) \in \overline{C}([0,T]; L^2(\Omega)) \mid \varphi(\cdot) \text{ has continuous paths } a.s.\},\\
&L^2(\Omega; C([0,T])) = \left\{\varphi(\cdot) \in C([0,T]; L^2(\Omega)) \,\middle|\, E\left[\sup_{t\in[0,T]} |\varphi(t)|^2\right] < \infty\right\}.
\end{aligned}$$

Note that process $\varphi(\cdot)$ belonging to the last four spaces above are not necessarily $\mathbb{F}$-adapted. Thus, naturally, we define

$$\begin{aligned}
&L^2_{\mathcal{F}}(0,T) = \{\varphi(\cdot) \in L^2((0,T)\times\Omega) \mid \varphi(\cdot) \text{ is } \mathbb{F}\text{ -adapted}\},\\
&C_{\mathcal{F}}([0,T]; L^2(\Omega)) = \{\varphi(\cdot) \in C([0,T]; L^2(\Omega)) \mid \varphi(\cdot) \text{ is } \mathbb{F}\text{-adapted}\},\\
&\overline{C}_{\mathcal{F}}([0,T]; L^2(\Omega)) = \{\varphi(\cdot) \in \overline{C}([0,T]; L^2(\Omega)) \mid \varphi(\cdot) \text{ is } \mathbb{F}\text{-adapted}\},\\
&L^2_{\mathcal{F}}(\Omega; C([0,T])) = \{\varphi(\cdot) \in L^2(\Omega; C([0,T])) \mid \varphi(\cdot) \text{ is } \mathbb{F}\text{-adapted}\}.
\end{aligned}$$

In the above definitions, we have suppressed the range space H in the notations. More carefully, one might want to use the notation $L^2(\Omega; H)$, $L^2((0,T)\times\Omega; H)$, and so on, to indicate the range space H. We will use such notations when the range space for the processes needs to be emphasized. We point out several facts about the spaces defined above. First of all, from the definition, the following chains of inclusions hold:

$$L^2(\Omega; C([0,T])) \subseteq C([0,T]; L^2(0,T)) \subseteq \overline{C}([0,T]; L^2(0,T)) \subseteq L^2((0,T)\times\Omega), \tag{2.1}$$

and

$$L^2_{\mathcal{F}}(\Omega; C([0,T])) \subseteq C_{\mathcal{F}}([0,T]; L^2(0,T)) \subseteq \overline{C}_{\mathcal{F}}([0,T]; L^2(0,T)) \subseteq L^2_{\mathcal{F}}(0,T). \tag{2.2}$$

Second, for any $\varphi(\cdot) \in \overline{C}([0,T]; L^2(\Omega))$, we only have the continuity of $t \mapsto \varphi(t)$ as a map from $[0,T]$ to $L^2(\Omega)$, and $\varphi(\cdot)$ does not necessarily have continuous paths. Actually, space $C([0,T]; L^2(\Omega))$ is not complete under the norm

$$\|\varphi(\cdot)\|_{\overline{C}([0,T];L^2(\Omega))} \triangleq \left\{\sup_{t\in[0,T]} E|\varphi(t)|^2\right\}^{\frac{1}{2}}, \tag{2.3}$$

and $\overline{C}([0,T]; L^2(\Omega))$ is the completion of $C([0,T]; L^2(\Omega))$ under norm (2.3).

Another important space that we are going to use is $L^2(0,T; L^2_{\mathcal{F}}(0,T))$. By definition, any process $Z : [0,T]^2\times\Omega \to H$ belongs to $L^2(0,T; L^2_{\mathcal{F}}(0,T))$ if it is $\mathcal{B}([0,T]^2)\otimes\mathcal{F}_T$-measurable;

for almost all $t \in [0, T]$, $Z(t, \cdot)$ is $\mathbb{F}$-adapted; and

$$E \int_0^T \int_0^T |Z(t,s)|^2 \, \mathrm{d}s \, \mathrm{d}t < \infty. \tag{2.4}$$

Similarly, we can define space $\overline{C}([0, T]; L^2_{\mathcal{F}}(0, T))$, etc.

Now, we look at the following integral equation:

$$Y(t) = \varphi(t) - \int_t^T Z(t,s) \, \mathrm{d}W(s), \quad t \in [0, T], \tag{2.5}$$

where $\varphi(\cdot) \in L^2((0, T) \times \Omega; \mathbb{R}^m)$ is given, and we are looking for a pair $(Y(\cdot), Z(\cdot\,,\cdot))$ of processes. Remember that, by definition, for almost all $t \in [0, T]$, $\varphi(t)$ is merely $\mathcal{F}_T$-measurable. Thus, $\varphi(\cdot)$ is not necessarily $\mathbb{F}$-adapted. Before going further, we introduce the following definition.

Definition 2.1. Any pair of stochastic processes $(Y(\cdot), Z(\cdot\,,\cdot)) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^m) \times L^2(0, T; L^2_{\mathcal{F}}(0, T; \mathbb{R}^{m\times d}))$ satisfying (2.5) is called an adapted solution to (2.5).

Clearly, (2.5) is a special case of (1.6). From the above definition, if $(Y(\cdot), Z(\cdot\,,\cdot))$ is an adapted solution to BSVIE (2.5), $Y(\cdot)$ is required to be $\mathbb{F}$-adapted and for almost all $t \in [0, T]$, $Z(t, \cdot)$ is required to be $\mathbb{F}$-adapted. We have the the following result concerning BSVIE (2.5).

Theorem 2.2. *For any $\varphi(\cdot) \in L^2((0, T) \times \Omega; \mathbb{R}^m)$, BSVIE (2.5) admits a unique adapted solution $(Y(\cdot), Z(\cdot\,,\cdot)) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^m) \times L^2(0, T; L^2_{\mathcal{F}}(0, T; \mathbb{R}^{m\times d}))$, with the following relation:*

$$Y(t) = E\varphi(t) + \int_0^t Z(t,s) \, \mathrm{d}W(s), \quad t \in [0, T]. \tag{2.6}$$

Moreover, the following estimate holds:

$$E \int_0^T |Y(t)|^2 \, \mathrm{d}t + E \int_0^T \int_0^T |Z(t,s)|^2 \, \mathrm{d}s \, \mathrm{d}t \le CE \int_0^T |\varphi(t)|^2 \, \mathrm{d}t. \tag{2.7}$$

Hereafter, $C > 0$ represents a generic constant which can be different at different places. In the case $\varphi(\cdot) \in \overline{C}([0, T]; L^2(\Omega; \mathbb{R}^m))$, one has $(Y(\cdot), Z(\cdot\,,\cdot)) \in \overline{C}_{\mathcal{F}}([0, T]; L^2(\Omega; \mathbb{R}^m)) \times \overline{C}([0, T]; L^2_{\mathcal{F}}(0, T; \mathbb{R}^{m\times d}))$ and

$$\sup_{t\in[0,T]} E|Y(t)|^2 + \sup_{t\in[0,T]} E \int_0^T |Z(t,s)|^2 \, \mathrm{d}s \le C \sup_{t\in[0,T]} E|\varphi(t)|^2. \tag{2.8}$$

Further, if $\varphi(\cdot) \in L^2(\Omega; C([0, T]; \mathbb{R}^m))$, then $Y(\cdot) \in L^2_{\mathcal{F}}(\Omega; C([0, T]; \mathbb{R}^m))$, and

$$E \left\{ \sup_{t\in[0,T]} |Y(t)|^2 \right\} \le E \left\{ \sup_{t\in[0,T]} |\varphi(t)|^2 \right\}. \tag{2.9}$$

Proof. First, we let $\varphi(\cdot) \in \overline{C}([0, T]; L^2(\Omega; \mathbb{R}^m))$. For any given $t \in [0, T]$, since $\varphi(t) \in L^2(\Omega; \mathbb{R}^m)$, by the martingale representation theorem [10,11], there exists a unique $Z(t, \cdot) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^{m\times d})$ (depending on the parameter $t \in [0, T]$) such that

$$\varphi(t) - E\varphi(t) = \int_0^T Z(t,s) \, \mathrm{d}W(s). \tag{2.10}$$

From this, we can easily obtain

$$\sup_{t\in[0,T]}\left\{E\int_0^T |Z(t,s)|^2\,\mathrm{d}s\right\} \le 2\sup_{t\in[0,T]} E\{|\varphi(t)|^2\}. \tag{2.11}$$

Further, by the continuity of $t\mapsto\varphi(t)$ (as a map from $[0,T]$ to $L^2(\Omega)$), we have

$$E\int_0^T |Z(t,s)-Z(t',s)|^2\,\mathrm{d}s \le 2E|\varphi(t)-\varphi(t')|^2 \to 0, \tag{2.12}$$

as $|t-t'|\to 0$. This means that $t\mapsto Z(t,\cdot)$ is continuous from $[0,T]$ to $L^2_{\mathcal{F}}(0,T;\mathbb{R}^{m\times d})$. Hence, together with (2.10), we see that $Z(\cdot\,,\cdot)\in\overline{C}([0,T];L^2_{\mathcal{F}}(0,T;\mathbb{R}^{m\times d}))$. Now, we define $Y(\cdot)$ by (2.6). Then $Y(\cdot)$ is $\mathbb{F}$-adapted, and combining with (2.10), we see that $(Y(\cdot),Z(\cdot\,,\cdot))$ is an adapted solution of (2.5). Taking conditional expectation in (2.5), one has

$$Y(t)=E[\varphi(t)|\mathcal{F}_t],\qquad t\in[0,T]. \tag{2.13}$$

Consequently,

$$\sup_{t\in[0,T]} E|Y(t)|^2 \le \sup_{t\in[0,T]} E|\varphi(t)|^2. \tag{2.14}$$

Moreover, for any $0\le t\le t'\le T$, using (2.6), we have

$$\begin{aligned} E|Y(t)-Y(t')|^2 &\le C\left\{E|\varphi(t)-\varphi(t')|^2 + E\int_0^T |Z(t,s)-Z(t',s)|^2\,\mathrm{d}s\right.\\ &\qquad\left. + E\int_0^T \mathbf{1}_{[t,t']}(s)|Z(t',s)|^2\,\mathrm{d}s\right\}\to 0, \end{aligned} \tag{2.15}$$

as $|t-t'|\to 0$. Thus, $Y(\cdot)\in\overline{C}_{\mathcal{F}}([0,T];L^2(\Omega;\mathbb{R}^m))$ and estimate (2.8) holds.

Next, from (2.10), we also see that

$$\int_0^T\int_0^T E|Z(t,s)|^2\,\mathrm{d}s\,\mathrm{d}t \le 2\int_0^T E|\varphi(t)|^2\,\mathrm{d}t. \tag{2.16}$$

Thus, for the general case that $\varphi(\cdot)\in L^2((0,T)\times\Omega;\mathbb{R}^m)$, we can use the approximation argument to obtain that (2.10) holds for some $Z(\cdot\,,\cdot)\in L^2(0,T;L^2_{\mathcal{F}}(0,T;\mathbb{R}^{m\times d}))$, for almost all $t\in[0,T]$. Then we again define $Y(\cdot)$ by (2.6) to get an adapted solution $(Y(\cdot),Z(\cdot\,,\cdot))$ of (2.5). In this case, $Y(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$. Estimate (2.7) then follows easily. The uniqueness of $(Y(\cdot),Z(\cdot\,,\cdot))$ follows immediately from estimate (2.7).

Finally, if $\varphi(\cdot)\in L^2(\Omega;C([0,T];\mathbb{R}^m))$, then from (2.13), we see that for any $t\to t'$,

$$\begin{aligned} |Y(t)-Y(t')| &= |E[\varphi(t)|\mathcal{F}_t]-E[\varphi(t')|\mathcal{F}_{t'}]|\\ &\le |E[\varphi(t)-\varphi(t')|\mathcal{F}_t]| + |E[\varphi(t')|\mathcal{F}_t]-E[\varphi(t')|\mathcal{F}_{t'}]|\to 0, \end{aligned} \tag{2.17}$$

proving the path continuity of $Y(\cdot)$. Estimate (2.9) follows easily from (2.13). □

Note that by (2.6), we also have

$$Y(t)=EY(t)+\int_0^t Z(t,s)\,\mathrm{d}W(s),\qquad t\in[0,T],\ a.s. \tag{2.18}$$

Finally, we emphasize that process $\varphi(\cdot)$ appearing in BSVIE (2.5) can be very general. One interesting case is as follows:

$$\varphi(t) = \int_t^T h(t,s)\,\mathrm{d}s, \quad t \in [0,T], \tag{2.19}$$

with $h(\cdot\,,\cdot) \in L^2((0,T)^2 \times \Omega; \mathbb{R}^m)$. We will see such a case in the next section.

3. Well-posedness for general BSVIEs

In this section, we consider the well-posedness of general BSVIE (1.6) with $h : [0,T] \times [0,T] \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times \mathbb{R}^{m\times d} \times \Omega \to \mathbb{R}^m$. Similar to Definition 2.1, we have the following.

Definition 3.1. A pair of processes $(Y(\cdot), Z(\cdot\,,\cdot)) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m) \times L^2(0,T; L^2_{\mathcal{F}}(0,T;\mathbb{R}^{m\times d}))$ satisfying (1.6) is called an adapted solution of (1.6).

We introduce the following standing assumption.

(H1) Let $h : [0,T] \times [0,T] \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times \mathbb{R}^{m\times d} \times \Omega \to \mathbb{R}^m$ satisfy the following:

(i) It is $\mathcal{B}([0,T]^2 \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times \mathbb{R}^{m\times d}) \otimes \mathcal{F}_T$-measurable.

(ii) There exists a constant $L > 0$ such that

$$\begin{cases} |h(t,s,0,0,0)| \le L, \\ |h(t,s,y,z,\zeta) - h(t,s,\bar{y},\bar{z},\bar{\zeta})| \le L(|y-\bar{y}| + |z-\bar{z}| + |\zeta-\bar{\zeta}|), \\ \quad \forall (t,s) \in [0,T]\times[0,T],\ y,\bar{y} \in \mathbb{R}^m,\ z,\bar{z},\zeta,\bar{\zeta} \in \mathbb{R}^{m\times d},\ a.s. \end{cases} \tag{3.1}$$

The following is an existence and uniqueness result for adapted solutions to BSVIE (1.6).

Theorem 3.2. *Suppose* (H1) *holds. Then for any $f(\cdot) \in L^2((0,T)\times\Omega;\mathbb{R}^m)$, BSVIE* (1.6) *admits a unique adapted solution $(Y(\cdot), Z(\cdot\,,\cdot)) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m) \times L^2(0,T; L^2_{\mathcal{F}}(0,T;\mathbb{R}^{m\times d}))$. Moreover, the following estimate holds:*

$$E\int_0^T |Y(t)|^2\,\mathrm{d}t + E\int_0^T\int_0^T |Z(t,s)|^2\,\mathrm{d}s\,\mathrm{d}t \le C\left(1 + E\int_0^T |f(t)|^2\,\mathrm{d}t\right). \tag{3.2}$$

If h is independent of ζ, $f(\cdot) \in \overline{C}([0,T]; L^2(\Omega;\mathbb{R}^m))$, one has $(Y(\cdot), Z(\cdot\,,\cdot)) \in \overline{C}_{\mathcal{F}}([0,T]; L^2(\Omega;\mathbb{R}^m)) \times \overline{C}([0,T]; L_{\mathcal{F}}(0,T;\mathbb{R}^{m\times d}))$, and

$$\sup_{t\in[0,T]} E|Y(t)|^2 + \sup_{t\in[0,T]} E\int_0^T |Z(t,s)|^2\,\mathrm{d}s \le C\left(1 + \sup_{t\in[0,T]} E|f(t)|^2\right). \tag{3.3}$$

Further, if $f(\cdot) \in L^2(\Omega; C([0,T];\mathbb{R}^m))$, then $Y(\cdot) \in L^2_{\mathcal{F}}(\Omega; C([0,T];\mathbb{R}^m))$ and

$$E\sup_{t\in[0,T]} |Y(t)|^2 \le C\left(1 + E\sup_{t\in[0,T]} |f(t)|^2\right). \tag{3.4}$$

Proof. For any $(y(\cdot), z(\cdot\,,\cdot))$ taken from $L^2_{\mathcal{F}}(0,T;\mathbb{R}^m) \times L^2(0,T; L^2_{\mathcal{F}}(0,T;\mathbb{R}^{m\times d}))$, let

$$\varphi(t) = f(t) + \int_t^T h(t,s,y(s),z(t,s),z(s,t))\,\mathrm{d}s, \quad t \in [0,T]. \tag{3.5}$$

Then by (H1),

$$E\int_0^T |\varphi(t)|^2\,\mathrm{d}t \le CE\left\{\int_0^T |f(t)|^2\,\mathrm{d}t + \int_0^T \left|\int_t^T h(t,s,y(s),z(t,s),z(s,t))\,\mathrm{d}s\right|^2 \mathrm{d}t\right\}$$
$$\le CE\left\{1 + \int_0^T |f(t)|^2\,\mathrm{d}t + \int_0^T |y(s)|^2\,\mathrm{d}s + \int_0^T\int_0^T |z(t,s)|^2\,\mathrm{d}s\,\mathrm{d}t\right\}. \tag{3.6}$$

Thus, $\varphi(\cdot) \in L^2((0,T)\times\Omega;\mathbb{R}^m)$. By Theorem 2.2, there exists a unique adapted solution $(Y(\cdot), Z(\cdot\,,\cdot))$ to the following BSVIE:

$$Y(t) = f(t) - \int_t^T h(t,s,y(s),z(t,s),z(s,t))\,\mathrm{d}s - \int_t^T \langle Z(t,s),\, \mathrm{d}W(s)\rangle. \tag{3.7}$$

Now, take another pair $(\bar y(\cdot), \bar z(\cdot\,,\cdot)) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)\times L^2(0,T;L^2_{\mathcal{F}}(0,T;\mathbb{R}^{m\times d}))$, and let $(\bar Y(\cdot), \bar Z(\cdot\,,\cdot))$ be the adapted solution of (3.7) with $(y(\cdot), z(\cdot\,,\cdot))$ replaced by $(\bar y(\cdot), \bar z(\cdot\,,\cdot))$. Then

$$[Y(t)-\bar Y(t)] = -\int_t^T \{h(t,s,y(s),z(t,s),z(s,t)) - h(t,s,\bar y(s),\bar z(t,s),\bar z(s,t))\}\,\mathrm{d}s$$
$$-\int_t^T [Z(t,s)-\bar Z(t,s)]\,\mathrm{d}W(s). \tag{3.8}$$

Hence, applying Theorem 2.2 on $[r,T]$, and (H1), we obtain

$$E\int_r^T |Y(t)-\bar Y(t)|^2\,\mathrm{d}t + E\int_r^T\int_r^T |Z(t,s)-\bar Z(t,s)|^2\,\mathrm{d}s\,\mathrm{d}t$$
$$\le CE\int_r^T \left|\int_r^T [h(t,s,y(s),z(t,s),z(s,t)) - h(t,s,\bar y(s),\bar z(t,s),\bar z(s,t))]\,\mathrm{d}s\right|^2 \mathrm{d}t$$
$$\le C(T-r)^2 E\int_r^T |y(t)-\bar y(t)|^2\,\mathrm{d}t + C(T-r)E\int_r^T\int_r^T |z(t,s)-\bar z(t,s)|^2\,\mathrm{d}s\,\mathrm{d}t. \tag{3.9}$$

Therefore the contraction mapping theorem applies to obtain a unique adapted solution $(Y(\cdot), Z(\cdot\,,\cdot)) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)\times L^2(0,T;L^2_{\mathcal{F}}(0,T;\mathbb{R}^{m\times d}))$. The remaining conclusions can be proved easily. $\square$

By a similar proof as the above, we are able to prove the following stability result for the adapted solutions to BSVIEs.

Theorem 3.3. *Let* (H1) *hold and* $f(\cdot), \bar f(\cdot) \in L^2((0,T)\times\Omega;\mathbb{R}^m)$. *Let* $(Y(\cdot), Z(\cdot\,,\cdot))$ *and* $(\bar Y(\cdot), \bar Z(\cdot\,,\cdot))$ *be corresponding adapted solutions to BSVIE* (3.1). *Then*

$$E\int_0^T |Y(t)-\bar Y(t)|^2\,\mathrm{d}t + E\int_0^T\int_0^T |Z(t,s)-\bar Z(t,s)|^2\,\mathrm{d}s\,\mathrm{d}t$$
$$\le CE\int_0^T |f(t)-\bar f(t)|^2\,\mathrm{d}t. \tag{3.10}$$

In the case that $f(\cdot), \bar{f}(\cdot) \in \overline{C}([0,T]; L^2(\Omega; \mathbb{R}^m))$, *one has*

$$\sup_{t\in[0,T]} E|Y(t)-\bar{Y}(t)|^2 + \sup_{t\in[0,T]} E\int_0^T |Z(t,s)-\bar{Z}(t,s)|^2\,ds \leq C \sup_{t\in[0,T]} E|f(t)-\bar{f}(t)|^2. \tag{3.11}$$

Further, if $f(\cdot), \bar{f}(\cdot) \in L^2(\Omega; C([0,T]; \mathbb{R}^m))$, *then*

$$E \sup_{t\in[0,T]} |Y(t)-\bar{Y}(t)|^2 \leq CE \sup_{t\in[0,T]} |f(t)-\bar{f}(t)|^2. \tag{3.12}$$

4. Duality principle

In this section, we discuss the duality relation between linear FSVIEs and linear BSVIEs. Such a relation plays an important role in some problems. We will see some applications a little later in this paper.

We consider the following linear FSVIE:

$$X(t) = g(t) + \int_0^t A_0(t,s)X(s)\,ds + \sum_{i=1}^d \int_0^t A_i(t,s)X(s)\,dW_i(s), \quad t\in[0,T], \tag{4.1}$$

where we assume that

$$A_i(\cdot\,,\cdot) \in L^\infty(0,T; L^\infty_{\mathcal{F}}(0,T;\mathbb{R}^{n\times n})), \quad 0\leq i\leq d. \tag{4.2}$$

Similar to the definition of $L^2(0,T; L^2_{\mathcal{F}}(0,T))$ in Section 2, (4.2) means that $A_i : [0,T]^2\times\Omega \to \mathbb{R}^{n\times n}$ is bounded, $\mathcal{B}([0,T]^2)\otimes\mathcal{F}_T$-measurable, and for almost all $t\in[0,T]$, $A_i(t,\cdot)$ is $\{\mathcal{F}_s\}_{s\geq 0}$-adapted. It is know that [1,5,13,17,19] for any $g(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$, FSVIE (4.1) admits a unique solution $X(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$. We have the following duality principle.

Theorem 4.1. *Let* (4.2) *hold and* $g(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$, $f(\cdot)\in L^2((0,T)\times\Omega;\mathbb{R}^n)$. *Let* $X(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ *be the solution of FSVIE* (4.1) *and* $(Y(\cdot), Z(\cdot\,,\cdot)) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)\times L^2(0,T; L^2_{\mathcal{F}}(0,T;\mathbb{R}^{n\times d}))$ *be the adapted solution to the following BSVIE:*

$$Y(t) = f(t) + \int_t^T \left[A_0(s,t)^T Y(s) + \sum_{i=1}^d A_i(s,t)^T Z_i(s,t)\right] ds - \int_t^T Z(t,s)\,dW(s), \quad t\in[0,T]. \tag{4.3}$$

Then the following relation holds:

$$E\int_0^T \langle Y(t), g(t)\rangle\,dt = E\int_0^T \langle f(t), X(t)\rangle\,dt. \tag{4.4}$$

BSVIE (4.3) is called the adjoint equation of FSVIE (4.1), and (4.4) is called the duality relation between FSVIE (4.1) and BSVIE (4.3). We point out that BSVIE (4.3) is of form (1.6). Also, we see that $f(\cdot)$ is not required to be $\{\mathcal{F}_r\}_{r\geq 0}$-adapted.

Proof. Let $(Y(\cdot), Z(\cdot\,,\cdot))$ be the adapted solution of (4.3). If we denote

$$\varphi(t) = f(t) + \int_t^T \left[A_0(s,t)^T Y(s) + \sum_{i=1}^d A_i(s,t)^T Z_i(s,t) \right] \mathrm{d}s, \tag{4.5}$$

then $(Y(\cdot), Z(\cdot\,,\cdot))$ is the (unique) adapted solution to the BSVIE:

$$Y(t) = \varphi(t) - \int_t^T Z(t,s)\,\mathrm{d}W(s), \quad t \in [0,T]. \tag{4.6}$$

By Theorem 2.2, it is necessary that

$$Y(t) = E\varphi(t) + \int_0^t Z(t,s)\,\mathrm{d}W(s), \quad t \in [0,T]. \tag{4.7}$$

Now, we observe the following:

$$\begin{aligned}
&E\int_0^T \langle Y(t), g(t)\rangle\,\mathrm{d}t \\
&\quad = E\int_0^T \left\langle Y(t), X(t) - \int_0^t A_0(t,s)X(s)\,\mathrm{d}s - \sum_{i=1}^d \int_0^t A_i(t,s)X(s)\,\mathrm{d}W_i(s) \right\rangle \mathrm{d}t \\
&\quad = E\int_0^T \langle Y(t), X(t)\rangle\,\mathrm{d}t - E\int_0^T\int_s^T \langle A_0(t,s)^T Y(t), X(s)\rangle\,\mathrm{d}t\,\mathrm{d}s \\
&\qquad - \sum_{i=1}^d E\int_0^T \left\langle E\varphi(t) + \int_0^t Z(t,s)\,\mathrm{d}W(s), \int_0^t A_i(t,s)X(s)\,\mathrm{d}W_i(s) \right\rangle \mathrm{d}t \\
&\quad = E\int_0^T \langle Y(t), X(t)\rangle\,\mathrm{d}t - E\int_0^T\int_t^T \langle A_0(s,t)^T Y(s), X(t)\rangle\,\mathrm{d}s\,\mathrm{d}t \\
&\qquad - \sum_{i=1}^d E\int_0^T\int_0^t \langle Z_i(t,s), A_i(t,s)X(s)\rangle\,\mathrm{d}s\,\mathrm{d}t \\
&\quad = E\int_0^T \left\langle Y(t) - \int_t^T \left[A_0(s,t)^T Y(s) - \sum_{i=1}^d A_i(s,t)^T Z_i(s,t) \right] \mathrm{d}s, X(t) \right\rangle \mathrm{d}t \\
&\quad = E\int_0^T \left\langle f(t) - \int_t^T Z(t,s)\,\mathrm{d}W(s), X(t) \right\rangle \mathrm{d}t = E\int_0^T \langle f(t), X(t)\rangle\,\mathrm{d}t.
\end{aligned} \tag{4.8}$$

This proves (4.4). □

Next, let us define a linear operator $\mathcal{A} : L^2_{\mathcal{F}}(0,T;\mathbb{R}^n) \to L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ as follows:

$$\begin{aligned}
(\mathcal{A}X)(t) &= \int_0^t A_0(t,s)X(s)\,\mathrm{d}s + \int_0^t \sum_{i=1}^d A_i(t,s)X(s)\,\mathrm{d}W_i(s), \\
&\quad \forall X(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n).
\end{aligned} \tag{4.9}$$

Under condition (4.2), we know that $\mathcal{A}$ is a bounded linear operator from the Hilbert space $L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ into itself. Thus, the adjoint $\mathcal{A}^*$ of $\mathcal{A}$ is well-defined. The idea of proving Theorem 4.1 can be used to prove the following result which gives an identification of $\mathcal{A}^*$.

Theorem 4.2. *Suppose (4.2) holds and $\mathcal{A}$ is defined by (4.9). Then for any $\eta(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$,*

$$(\mathcal{A}^*\eta)(t) = Y(t), \quad t \in [0,T], \text{a.s.}, \tag{4.10}$$

where $(Y(\cdot), Z(\cdot\,,\cdot))$ is the unique adapted solution of the following BSVIE:

$$\begin{aligned} Y(t) &= \int_t^T \left[A_0(s,t)^T \eta(s) - \sum_{i=1}^d A_i(s,t)^T Z_i(s,t) \right] \mathrm{d}s - \int_t^T Z(t,s)\,\mathrm{d}W(s), \\ &\quad t \in [0,T]. \end{aligned} \tag{4.11}$$

Proof. For any $\eta(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$, by Theorem 2.2, we know that BSVIE (4.11) admits a unique adapted solution $(Y(\cdot), Z(\cdot\,,\cdot)) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n) \times L^2(0,T;L^2_{\mathcal{F}}(0,T;\mathbb{R}^n))$. If we call

$$\begin{aligned} \varphi(t) &= Y(t) + \eta(t) - \int_t^T \left[A_0(s,t)^T \eta(s) - \sum_{i=1}^d A_i(s,t)^T Z_i(s,t) \right] \mathrm{d}s \\ &\equiv \eta(t) - \int_t^T Z(t,s)\,\mathrm{d}W(s), \quad t \in [0,T], \end{aligned} \tag{4.12}$$

then $(\eta(\cdot), Z(\cdot\,,\cdot))$ is the unique adapted solution of the following BSVIE:

$$\eta(t) = \varphi(t) + \int_t^T Z(t,s)\,\mathrm{d}W(s), \quad t \in [0,T]. \tag{4.13}$$

Thus, by Theorem 2.2, we must have

$$\eta(t) = E\varphi(t) - \int_0^t Z(t,s)\,\mathrm{d}W(s), \quad t \in]0,T]. \tag{4.14}$$

Consequently, it follows that

$$\begin{aligned} &E\int_0^T \langle (\mathcal{A}^*\eta)(t), X(t)\rangle\,\mathrm{d}t \equiv E\int_0^T \langle \eta(t), (\mathcal{A}X)(t)\rangle\,\mathrm{d}t \\ &= E\int_0^T \left\{ \left\langle \eta(t), \int_0^t A_0(t,s)X(s)\,\mathrm{d}s + \sum_{i=1}^d \int_0^t A_i(t,s)X(s)\,\mathrm{d}W_i(s) \right\rangle \right\} \mathrm{d}t \\ &= E\int_0^T \int_s^T \langle A_0(t,s)^T \eta(t), X(s)\rangle\,\mathrm{d}t\,\mathrm{d}s \\ &\quad + E\int_0^T \left\langle E\varphi(t) - \int_0^t Z(t,s)\,\mathrm{d}W(s), \sum_{i=1}^d \int_0^t A_i(t,s)X(s)\,\mathrm{d}W_i(s) \right\rangle \mathrm{d}t \\ &= E\int_0^T \int_t^T \langle A_0(s,t)^T \eta(s), X(t)\rangle\,\mathrm{d}s\,\mathrm{d}t \\ &\quad - \sum_{i=1}^d E\int_0^T \int_0^t \langle Z_i(t,s), A_i(t,s)X(s)\rangle\,\mathrm{d}s\,\mathrm{d}t \\ &= E\int_0^T \left\langle \int_t^T \left[A_0(s,t)^T \eta(s) - \sum_{i=1}^d A_i(s,t)^T Z_i(s,t) \right] \mathrm{d}s, X(t) \right\rangle \mathrm{d}t \\ &= E\int_0^T \langle Y(t), X(t)\rangle\,\mathrm{d}t. \end{aligned} \tag{4.15}$$

Since $X(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ is arbitrary, and $Y(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$, we obtain (4.10). □

From (4.10) and (4.11), we also have

$$(\mathcal{A}^*\eta)(t) = E\left[\int_t^T \left[A_0(s,t)^T\eta(s) + \sum_{i=1}^d A_i(s,t)^T Z_i(s,t)\right] \mathrm{d}s \,\middle|\, \mathcal{F}_t\right], \quad t \in [0,T], \tag{4.16}$$

with $Z(\cdot\,,\cdot)$ uniquely determined by (4.14) which is equivalent to

$$\eta(t) = E\eta(t) - \int_0^t Z(t,s)\,\mathrm{d}W(s), \quad t \in [0,T]. \tag{4.17}$$

5. A comparison theorem

In this section, we are going to establish a comparison theorem for adapted solutions to BSVIEs.

First, let us consider the following FSVIE:

$$X(t) = g(t) + \int_0^t a(t,s)X(s)\,\mathrm{d}s + \int_0^t \langle b(t,s)X(s), \mathrm{d}W(s)\rangle, \quad t \in [0,T]. \tag{5.1}$$

Here, $n = 1$ and $d \geq 1$. We have the following result.

Lemma 5.1. *Let $a(\cdot\,,\cdot) \in C([0,T]; L^\infty_{\mathcal{F}}(0,T;\mathbb{R}))$ and $b(\cdot\,,\cdot) \in C([0,T]; L^\infty_{\mathcal{F}}(0,T;\mathbb{R}^d))$. Then for any $g(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R})$, with $g(t) \geq 0$, the solution $X(\cdot)$ of (5.1) satisfies*

$$X(t) \geq 0, \quad t \in [0,T], \text{ a.s.} \tag{5.2}$$

The above result should exist in some literature. Since we could not find an exact reference, for the reader's convenience, a proof is presented here.

Proof. First, we let

$$\begin{cases} a(t,s) = \sum_{i\geq 0} a_i(s)\mathbf{1}_{[\tau_i,\tau_{i+1})}(t), \quad b(t,s) = \sum_{i\geq 0} b_i(s)\mathbf{1}_{[\tau_i,\tau_{i+1})}(t), \\ g(t) = \sum_{i\geq 0} g_i \mathbf{1}_{[\tau_i,\tau_{i+1})}(t), \end{cases} \tag{5.3}$$

with $0 = \tau_0 < \tau_1 < \cdots$ being a sequence of $\mathbb{F}$-stopping times, $a_i(\cdot)$ and $b_i(\cdot)$ being some $\mathbb{F}$-adapted and bounded processes, and each $g_i \geq \delta > 0$ is $\mathcal{F}_{\tau_i}$-measurable. Then on $[0, \tau_1)$, (5.1) is equivalent to

$$X(t) = g_0 + \int_0^t a_0(s)X(s)\,\mathrm{d}s + \int_0^t \langle b_0(s)X(s),\ \mathrm{d}W(s)\rangle, \tag{5.4}$$

which is an FSDE. Thus, by a well-known comparison theorem [8,15], we obtain

$$X(t) > 0, \quad t \in [0,\tau_1], \text{ a.s.} \tag{5.5}$$

By induction on the intervals $[\tau_i, \tau_{i+1}]$, we see that (5.2) holds for the current special case. Then by approximation, we obtain the general case easily. □

Having the above result, we would like to prove a comparison theorem for BSVIEs. Recall that for BSDEs, a comparison theorem can be proved by using Itô's formula [8,15]. However, for

BSVIEs, it is not clear if one can use a similar technique. Fortunately, we have proved a duality principle in the previous section which can be used to prove a comparison theorem for certain BSVIEs. Due the nature of the duality principle, instead of BSVIE (1.6), we have to consider the following BSVIE:

$$Y(t) = f(t) - \int_t^T h(t, s, Y(s), Z(s, t))\,\mathrm{d}s - \int_t^T \langle Z(t, s), \mathrm{d}W(s)\rangle, \quad t \in [0, T]. \tag{5.6}$$

Here $h : [0, T] \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \times \Omega \to \mathbb{R}$. Note that (5.6) is not the general BSDE (1.6). It is not clear at this moment if BSVIE (1.6) has a similar comparison result. Now, we state and prove the following comparison theorem.

Theorem 5.2. *Let $h, \bar{h} : [0, T] \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ satisfy* (H1)*, and let $f(\cdot), \bar{f}(\cdot) \in L^2((0, T) \times \Omega; \mathbb{R})$ such that*

$$h(t, s, y, z) \le \bar{h}(t, s, y, z), \quad \forall (t, s, y, z) \in [0, T] \times [0, T] \times \mathbb{R} \times \mathbb{R}^d, a.s., \tag{5.7}$$

and

$$f(t) \ge \bar{f}(t), \quad t \in [0, T], a.s. \tag{5.8}$$

Let $(Y(\cdot), Z(\cdot\,, \cdot))$ be the adapted solution of BSVIE (5.6), and $(\bar{Y}(\cdot), \bar{Z}(\cdot\,, \cdot))$ be the adapted solution of BSVIE (5.6) with $h(\cdot)$ and $f(\cdot)$ replaced by $\bar{h}(\cdot)$ and $\bar{f}(\cdot)$, respectively. Then the following holds:

$$Y(t) \ge \bar{Y}(t), \quad \forall t \in [0, T], a.s. \tag{5.9}$$

Proof. By (5.6), we have

$$\begin{aligned} &Y(t) - \bar{Y}(t) \\ &\quad = f(t) - \bar{f}(t) - \int_t^T [h(t, s, Y(s), Z(s, t)) - \bar{h}(t, s, \bar{Y}(s), \bar{Z}(s, t))]\,\mathrm{d}s \\ &\qquad - \int_t^T [Z(t, s) - \bar{Z}(t, s)]\,\mathrm{d}W(s) \\ &\quad \equiv \varphi(t) - \int_t^T \{a(s, t)[Y(s) - \bar{Y}(s)] + \langle b(s, t), Z(s, t) - \bar{Z}(s, t)\rangle\}\,\mathrm{d}s \\ &\qquad - \int_t^T \langle Z(t, s) - \bar{Z}(t, s),\ \mathrm{d}W(s)\rangle, \end{aligned} \tag{5.10}$$

where

$$\begin{cases} \varphi(t) = f(t) - \bar{f}(t) - \displaystyle\int_t^T [h(t, s, \bar{Y}(s), \bar{Z}(s, t)) - \bar{h}(t, s, \bar{Y}(s), \bar{Z}(s, t))]\,\mathrm{d}s \\ a(s, t) = \displaystyle\int_0^1 h_y(t, s, \bar{Y}(s) + \lambda[Y(s) - \bar{Y}(s)], \bar{Z}(s, t) + \lambda[Z(s, t) - \bar{Z}(s, t)])\,\mathrm{d}\lambda, \\ b(s, t) = \displaystyle\int_0^1 h_z(t, s, \bar{Y}(s) + \lambda[Y(s) - \bar{Y}(s)], \bar{Z}(s, t) + \lambda[Z(s, t) - \bar{Z}(s, t)])\,\mathrm{d}\lambda. \end{cases} \tag{5.11}$$

Here, we assume that h_y and h_z exists (the general case can be treated by approximation). Now, for any $g(\cdot) \in L^2_{\mathcal{F}}(0, T; \mathbb{R})$, non-negative valued, let $X(\cdot)$ be the solution of FSVIE (5.1) with

$a(\cdot\,,\cdot)$ and $b(\cdot\,,\cdot)$ given by (5.11). Then Lemma 5.1 tells us that $X(\cdot)$ is also non-negative valued. Hence, by Theorem 4.1, making use of conditions (5.7) and (5.8), we have

$$E\int_0^T [Y(t)-\bar{Y}(t)]g(t)\,\mathrm{d}t = E\int_0^T \varphi(t)X(t)\,\mathrm{d}t \geq 0. \tag{5.12}$$

Since $g(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R})$ is arbitrary, we see that (5.9) necessarily holds. $\square$

Note that in BSDEs, the comparison theorem is proved independent of the corresponding result for FSDEs and the duality principle is not necessary [8,15]. Here, however, both play crucial roles.

6. A maximum principle for optimal controls

In this section, we briefly present a Pontryagin's maximum principle for optimal controls of stochastic integral equations. This is actually one of the important motivations of introducing BSVIEs and establishing the duality principle between FSVIEs and BSVIEs. A systematic presentation of optimal stochastic control theory can be found in the book [23] (and the references cited therein).

To simplify the presentation, we only consider the one-dimensional case, i.e., $n=m=d=1$. The general problem of optimal control for stochastic integral equations will be much more involved and will be carried out in a forthcoming paper.

Consider the following controlled stochastic integral equation:

$$X(t)=x+\int_0^t b(t,s,X(s),u(s))\,\mathrm{d}s+\int_0^t \sigma(t,s,X(s),u(s))\,\mathrm{d}W(s),\quad t\in[0,T], \tag{6.1}$$

where $X(\cdot)$ and $u(\cdot)$ are state and control processes, respectively; $b,\sigma:[0,T]\times[0,T]\times\mathbb{R}\times U\to\mathbb{R}$ are given maps and $x\in\mathbb{R}$; U is a bounded interval in $\mathbb{R}$. The cost functional is defined to be the following Lagrange form:

$$J(u(\cdot))=E\int_0^T h(t,X(t),u(t))\,\mathrm{d}t, \tag{6.2}$$

with $h:[0,T]\times\mathbb{R}\times U\to\mathbb{R}$ being a given map as well. In the above, all the functions can be random, and initial state x can also be replaced by a stochastic process.

We now introduce the following assumption. The conditions assumed are more than enough. One can relax many of them. But we prefer these strong conditions to make the presentation simpler.

(H2) Let b, σ, and h be continuous in all of their arguments, and differentiable in the variables X and u, with bounded derivatives.

We let

$$\mathcal{U}\stackrel{\Delta}{=}\{u:[0,T]\times\Omega\to U\mid u(\cdot)\text{ is }\{\mathcal{F}_r\}_{r\geq 0}\text{-progressively measurable}\}. \tag{6.3}$$

It is not hard to show that [1,5,13,17,19] under (H2), for any $x\in\mathbb{R}$ and $u(\cdot)\in\mathcal{U}$, (6.1) admits a unique solution $X(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R})$. Thus the cost functional $J(u(\cdot))$ is well-defined. Our optimal control problem can be stated as follows.

Problem (C). Find a $\bar{u}(\cdot)\in\mathcal{U}$ such that

$$J(\bar{u}(\cdot))=\inf_{u(\cdot)\in\mathcal{U}} J(u(\cdot)). \tag{6.4}$$

Any $\bar{u}(\cdot)$ satisfying (6.4) is called an optimal control of Problem (C), the corresponding state process $\bar{X}(\cdot)$ is called an optimal state process and $(\bar{X}(\cdot), \bar{u}(\cdot))$ is called an optimal pair.

The main result of this section is the following Pontryagin type maximum principle.

Theorem 6.1. *Let* (H2) *hold. Let* $(\bar{X}(\cdot), \bar{u}(\cdot))$ *be an optimal pair of Problem* (C)*. Then there exists a unique adapted solution* $(Y(\cdot), Z(\cdot\,,\cdot))$ *of the following BSVIE*

$$\begin{cases} Y(t) = -h_x(t, \bar{X}(t), \bar{u}(t)) + \displaystyle\int_t^T [b_x(s, t, \bar{X}(t), \bar{u}(t))Y(s) \\ \qquad\qquad + \sigma_x(s, t, \bar{X}(t), \bar{u}(t))Z(s, t)]\,\mathrm{d}s - \displaystyle\int_t^T Z(t, s)\,\mathrm{d}W(s), \quad t \in [0, T] \end{cases} \tag{6.5}$$

such that

$$\begin{aligned} &\left\{ E\left[\int_t^T [b_u(s, t, \bar{X}(t), \bar{u}(t))Y(s) + \sigma_u(s, t, \bar{X}(t), \bar{u}(t))Z(s, t)]\,\mathrm{d}s \,\middle|\, \mathcal{F}_t \right] \right. \\ &\qquad \left. - h_u(t, \bar{X}(t), \bar{u}(t)) \right\} [u - \bar{u}(t)] \le 0, \quad \forall u \in U, t \in [0, T], a.s. \end{aligned} \tag{6.6}$$

Proof. Let $(\bar{X}(\cdot), \bar{u}(\cdot))$ be an optimal pair. Take any $u(\cdot) \in \mathcal{U}$. Since $\mathcal{U}$ is convex, for any $\varepsilon \in (0, 1)$,

$$u_\varepsilon(\cdot) \overset{\Delta}{=} \bar{u}(\cdot) + \varepsilon[u(\cdot) - \bar{u}(\cdot)] \in \mathcal{U}. \tag{6.7}$$

Let $X_\varepsilon(\cdot)$ be the solution of (6.1) corresponding to $u_\varepsilon(\cdot)$. Define

$$\xi_\varepsilon(t) = \frac{X_\varepsilon(t) - \bar{X}(t)}{\varepsilon}, \quad t \in [0, T]. \tag{6.8}$$

Then $\xi_\varepsilon(\cdot) \to \xi(\cdot)$ in $L^2_{\mathcal{F}}(0, T; \mathbb{R})$ with $\xi(\cdot)$ satisfying the following:

$$\begin{aligned} \xi(t) = &\int_0^t \{b_x(t, s, \bar{X}(s), \bar{u}(s))\xi(s) + b_u(t, s, \bar{X}(s), \bar{u}(s))[u(s) - \bar{u}(s)]\}\,\mathrm{d}s \\ &+ \int_0^t \{\sigma_x(t, s, \bar{X}(s), \bar{u}(s))\xi(s) + \sigma_u(t, s, \bar{X}(s), \bar{u}(s))[u(s) - \bar{u}(s)]\}\,\mathrm{d}W(s). \end{aligned} \tag{6.9}$$

Now, we let $(Y(\cdot), Z(\cdot\,,\cdot))$ be the unique adapted solution to BSVIE (6.5). By the optimality of $(\bar{X}(\cdot), \bar{u}(\cdot))$, and the duality principle (Theorem 4.1), we have

$$\begin{aligned} 0 \le\ &\frac{J(u_\varepsilon(\cdot)) - J(\bar{u}(\cdot))}{\varepsilon} \\ \to\ &E\int_0^T \{h_x(t, \bar{X}(t), \bar{u}(t))\xi(t) + h_u(t, \bar{X}(t), \bar{u}(t))[u(t) - \bar{u}(t)]\}\,\mathrm{d}t \\ =\ &-E\int_0^T Y(t) \left\{ \int_0^t b_u(t, s, \bar{X}(s), \bar{u}(s))[u(s) - \bar{u}(s)]\,\mathrm{d}s \right. \\ &\left. + \int_0^t \sigma_u(t, s, \bar{X}(s), \bar{u}(s))[u(s) - \bar{u}(s)]\,\mathrm{d}W(s) \right\} \mathrm{d}t \\ &+ E\int_0^T h_u(t, \bar{X}(t), \bar{u}(t))[u(t) - \bar{u}(t)]\,\mathrm{d}t \end{aligned}$$

$$= -E\int_0^T \left\{ \int_t^T [b_u(s,t,\bar X(t),\bar u(t))Y(s) + \sigma_u(s,t,\bar X(t),\bar u(t))Z(s,t)]\,\mathrm{d}s \right.$$
$$\left. - h_u(t,\bar X(t),\bar u(t)) \right\} [u(t) - \bar u(t)]\,\mathrm{d}t. \tag{6.10}$$

Then (6.6) follows. □

Note that if we define

$$H(t,\bar X(t),\bar u(t),Y(\cdot),Z(\cdot\,,t),u) = \left\{ E\left[\int_t^T [b_u(s,t,\bar X(t),\bar u(t))Y(s) \right.\right.$$
$$\left.\left. + \sigma_u(s,t,\bar X(t),\bar u(t))Z(s,t)]\,\mathrm{d}s \,\middle|\, \mathcal{F}_t \right] - h_u(t,\bar X(t),\bar u(t)) \right\} u, \tag{6.11}$$

then (6.6) can be written as

$$H(t,\bar X(t),\bar u(t),Y(\cdot),Z(\cdot\,,t),\bar u(t)) = \max_{u\in U} H(t,\bar X(t),\bar u(t),Y(\cdot),Z(\cdot\,,t),u). \tag{6.12}$$

We call $H(\cdot)$ defined by (6.11) the Hamiltonian of our optimal control problem, call (6.6) (and (6.12)) the maximum condition, and call (6.5) the adjoint equation of (6.1), along the optimal pair $(\bar X(\cdot),\bar u(\cdot))$.

Note that in the current case, U is an interval in $\mathbb{R}$. Suppose it is a closed interval. Then by scaling and shifting, we may assume that $U = [-1,1]$. Consequently, the maximum condition (6.6) is equivalent to the following:

$$\bar u(t) = \operatorname{sgn}\left\{ E\left[\int_t^T [b_u(s,t,\bar X(t),\bar u(t))Y(s) + \sigma_u(s,t,\bar X(t),\bar u(t))Z(s,t)]\,\mathrm{d}s \,\middle|\, \mathcal{F}_t \right] \right.$$
$$\left. - h_u(t,\bar X(t),\bar u(t)) \right\}, \quad t\in[0,T],\ a.s. \tag{6.13}$$

We may call (6.13) a "Stochastic Bang–Bang Principle" for our optimal control problem.

Finally, by putting (6.1), (6.5) and (6.6) together (dropping the bars in $(\bar X(\cdot),\bar u(\cdot))$), we obtain the following system:

$$\begin{cases} X(t) = x + \int_0^t b(t,s,X(s),u(s))\,\mathrm{d}s + \int_0^t \sigma(t,s,X(s),u(s))\,\mathrm{d}W(s), \\ Y(t) = -h_x(t,X(t),u(t)) + \int_t^T [b_x(s,t,X(t),u(t))Y(s) \\ \qquad + \sigma_x(s,t,X(t),u(t))Z(s,t)]\,\mathrm{d}s - \int_t^T Z(t,s)\,\mathrm{d}W(s), \\ \left\{ E\left[\int_t^T [b_u(s,t,X(t),u(t))Y(s) + \sigma_u(s,t,X(t),u(t))Z(s,t)]\,\mathrm{d}s \,\middle|\, \mathcal{F}_t \right] \right. \\ \qquad \left. - h_u(t,X(t),u(t)) \right\} [v - u(t)] \le 0, \quad \forall v\in U, t\in[0,T],\ a.s. \end{cases} \tag{6.14}$$

This is a couple of FSVIE and BSVIE systems. The coupling is through the maximum condition (via $u(\cdot)$). We call (6.14) a forward–backward stochastic integral equation (FBSVIE, for short). Such kinds of equations are still under careful investigation. We hope to publish some results for FBSVIEs in the near future.

Acknowledgement

The author would like to thank the anonymous referee for useful suggestions and comments which led to an improvement of Theorems 4.1 and 5.2.

References

[1] M. Berger, V. Mizel, Volterra equations with Itô integrals, I, II, J. Int. Eqs. 2 (1980) 187–245, 319–337.

[2] J.M. Bismut, Théorie Probabiliste du Contrôle des Diffusions, in: Memoir Amer. Math. Soc., vol. 176, Providence, RI, 1973.

[3] P. Briand, B. Delyon, Y. Hu, E. Pardoux, L. Stoica, L^p solutions of backward stochastic differential equations, Stochastic Process Appl. 108 (2003) 109–129.

[4] Z. Chen, L. Epstein, Ambiguity, risk and asset returns in continuous time, Econometrica 70 (2001) 1403–1444.

[5] C. Doléans-Dade, On the existence and unicity of solutions of the stochastic integral equations, Z. Wahrsch. verw. Gebiete 36 (1976) 93–101.

[6] D. Duffie, L. Epstein, Stochastic differential utility, Econometrica 60 (1992) 353–394.

[7] D. Duffie, C. Skiadas, Continuous-time security pricing: a utility gradient approach, J. Math. Econom. 23 (1994) 107–131.

[8] N. El Karoui, S. Peng, M.C. Quenez, Backward stochastic differential equations in finance, Math. Finance 7 (1997) 1–71.

[9] N. El Karoui, S. Peng, M.C. Quenez, A dynamic maximum principle for the optimization of recursive utilities under constraints, Ann. Appl. Probab. 11 (2001) 664–693.

[10] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes, 2nd ed., North-Holland, 1989.

[11] I. Karatzas, S.E. Shreve, Brownian Motion and Stochastic Calculus, Springer-Verlag, 1988.

[12] M. Kobylanski, Backward stochastic differential equations and partial differential equations with quadratic growth, Ann. Probab. 28 (2000) 558–602.

[13] A.M. Kolodii, On the existence of solutions of stochastic Volterra equation, Theory Random Proc. 11 (1983) 51–57 (in Russian).

[14] A. Lazrak, M.C. Quenez, A generalized stochastic utility, Math. Oper. Res. 28 (2003) 154–180.

[15] J. Ma, J. Yong, Forward–Backward Stochastic Differential Equations and Their Applications, Springer-Verlag, Berlin, 1999.

[16] E. Pardoux, S. Peng, Adapted solutions of backward stochastic equations, Systems Control Lett. 14 (1990) 55–61.

[17] E. Pardoux, P. Protter, Stochastic Volterra equations with anticipating coefficients, Ann. Probab. 18 (1990) 1635–1655.

[18] S. Peng, Nonlinear expectations, nonlinear evaluations and risk measures (preprint).

[19] P. Protter, Volterra equations driven by semimartingales, Ann. Probab. 13 (1985) 519–530.

[20] M. Schroder, C. Skiadas, Optimal consumption and portfolio selection with stochastic utility, J. Econom. Theory 89 (1999) 68–126.

[21] C. Skiadas, Robust control and recursive utility, Finance Stoch. 7 (2003) 475–489.

[22] Y. Shiota, A linear stochastic integral equation containing the extended Itô integral, Math. Rep. Toyama Univ. 9 (1986) 61–74.

[23] J. Yong, X.Y. Zhou, Stochastic Controls: Hamiltonian Systems and HJB Equations, Springer-Verlag, New York, 1999.